

SEQUENCE LISTING

:110> ARAKI, HIROYUKI

ENDO, KEIJI

HAGIHARA, HIROSHI

IGARASHI, KAZUAKI

HAYASHI, YASUHIRO

OZAKI, KATSUYA

<120> HIGHLY PRODUCTIVE ALPHA-AMYLASES

<130> 214377US0

<140> 09/971,611

<141> 2001-10-09

<150> JP 2000/310605

<151> 2001-10-11

<160> 51

<170> PatentIn version 3.1

<210> 1

<211> 1786

<212> DNA

<213> Bacillus sp. KSM-AP1378

<220>

<221> sig\_peptide

<222> (155)..(247)

<223>

<220>

<221> mat\_peptide

<222> (248)..()

<223>

<220>

<221> CDS

<222> (155)..(1702)

<223>

<400> 1

cagcgtgata atataaattt gaaatgaaca cctatgaaaa tatggtagcg attgcgcgac 60

gagaaaaaac ttgggagtta ggaagtgata ttaaaggatt ttttttgact tgttgtgaaa 120

acgcttgcat aaattgaagg agaggggtgct tttt atg aaa ctt cat aac cgt ata 175

Met Lys Leu His Asn Arg Ile  
-30 -25

att agc gta cta tta aca cta ttg tta gct gta gct gtt ttg ttt cca 223

Ile Ser Val Leu Leu Thr Leu Leu Leu Ala Val Ala Val Leu Phe Pro  
-20 -15 -10

tat atg acg gaa cca gca caa gcc cat cat aat ggg acg aat ggg acc 271

Tyr Met Thr Glu Pro Ala Gln Ala His His Asn Gly Thr Asn Gly Thr  
-5 -1 1 5

atg atg cag tat ttt gaa tgg cat ttg cca aat gac ggg aac cac tgg 319

Met Met Gln Tyr Phe Glu Trp His Leu Pro Asn Asp Gly Asn His Trp  
10 15 20

aac	agg	tta	cga	gat	gac	gca	gct	aac	tta	aag	agt	aaa	ggg	att	acc	367
Asn	Arg	Leu	Arg	Asp	Asp	Ala	Ala	Asn	Leu	Lys	Ser	Lys	Gly	Ile	Thr	
25					30					35					40	
gct	gtt	tgg	att	cct	cct	gca	tgg	aag	ggg	act	tcg	caa	aat	gat	gtt	415
Ala	Val	Trp	Ile	Pro	Pro	Ala	Trp	Lys	Gly	Thr	Ser	Gln	Asn	Asp	Val	
				45					50					55		
ggg	tat	ggt	gcc	tat	gat	ttg	tac	gat	ctt	ggt	gag	ttt	aac	caa	aag	463
Gly	Tyr	Gly	Ala	Tyr	Asp	Leu	Tyr	Asp	Leu	Gly	Glu	Phe	Asn	Gln	Lys	
			60					65					70			
gga	acc	gtc	cgt	aca	aaa	tat	ggc	aca	agg	agt	cag	ttg	caa	ggt	gcc	511
Gly	Thr	Val	Arg	Thr	Lys	Tyr	Gly	Thr	Arg	Ser	Gln	Leu	Gln	Gly	Ala	
		75					80					85				
gtg	aca	tct	ttg	aaa	aat	aac	ggg	att	caa	gtt	tat	ggg	gat	gtc	gtg	559
Val	Thr	Ser	Leu	Lys	Asn	Asn	Gly	Ile	Gln	Val	Tyr	Gly	Asp	Val	Val	
	90					95					100					
atg	aat	cat	aaa	ggt	gga	gca	gac	ggg	aca	gag	atg	gta	aat	gcg	gtg	607
Met	Asn	His	Lys	Gly	Gly	Ala	Asp	Gly	Thr	Glu	Met	Val	Asn	Ala	Val	
105					110					115					120	
gaa	gtg	aac	cga	agc	aac	cga	aac	caa	gaa	ata	tca	ggt	gaa	tac	acc	655
Glu	Val	Asn	Arg	Ser	Asn	Arg	Asn	Gln	Glu	Ile	Ser	Gly	Glu	Tyr	Thr	
				125					130					135		
att	gaa	gca	tgg	acg	aaa	ttt	gat	ttc	cct	gga	aga	gga	aat	acc	cat	703
Ile	Glu	Ala	Trp	Thr	Lys	Phe	Asp	Phe	Pro	Gly	Arg	Gly	Asn	Thr	His	
			140					145					150			
tcc	aac	ttt	aaa	tgg	cgc	tgg	tat	cat	ttt	gat	ggg	aca	gat	tgg	gat	751
Ser	Asn	Phe	Lys	Trp	Arg	Trp	Tyr	His	Phe	Asp	Gly	Thr	Asp	Trp	Asp	
		155					160					165				
cag	tca	cgt	cag	ctt	cag	aac	aaa	ata	tat	aaa	ttc	aga	ggt	acc	gga	799
Gln	Ser	Arg	Gln	Leu	Gln	Asn	Lys	Ile	Tyr	Lys	Phe	Arg	Gly	Thr	Gly	
		170				175					180					
aag	gca	tgg	gac	tgg	gaa	gta	gat	ata	gag	aac	ggc	aac	tat	gat	tac	847
Lys	Ala	Trp	Asp	Trp	Glu	Val	Asp	Ile	Glu	Asn	Gly	Asn	Tyr	Asp	Tyr	
185					190					195					200	
ctt	atg	tat	gca	gac	att	gat	atg	gat	cat	cca	gaa	gta	atc	aat	gaa	895
Leu	Met	Tyr	Ala	Asp	Ile	Asp	Met	Asp	His	Pro	Glu	Val	Ile	Asn	Glu	
				205					210					215		
ctt	aga	aat	tgg	gga	gtt	tgg	tat	aca	aat	aca	ctt	aat	cta	gat	gga	943
Leu	Arg	Asn	Trp	Gly	Val	Trp	Tyr	Thr	Asn	Thr	Leu	Asn	Leu	Asp	Gly	
			220					225					230			

ttt	aga	atc	gat	gct	gtg	aaa	cat	att	aaa	tac	agc	tat	acg	aga	gat	991
Phe	Arg	Ile	Asp	Ala	Val	Lys	His	Ile	Lys	Tyr	Ser	Tyr	Thr	Arg	Asp	
		235					240					245				
tgg	cta	aca	cat	gtg	cgt	aac	acc	aca	ggc	aaa	cca	atg	ttt	gca	gtt	1039
Trp	Leu	Thr	His	Val	Arg	Asn	Thr	Thr	Gly	Lys	Pro	Met	Phe	Ala	Val	
	250					255					260					
gca	gaa	ttt	tgg	aaa	aat	gac	ctt	gct	gca	atc	gaa	aac	tat	tta	aat	1087
Ala	Glu	Phe	Trp	Lys	Asn	Asp	Leu	Ala	Ala	Ile	Glu	Asn	Tyr	Leu	Asn	
265					270					275					280	
aaa	aca	agt	tgg	aat	cac	tcc	gtg	ttc	gat	gtt	cct	ctt	cat	tat	aat	1135
Lys	Thr	Ser	Trp	Asn	His	Ser	Val	Phe	Asp	Val	Pro	Leu	His	Tyr	Asn	
				285					290					295		
ttg	tac	aat	gca	tct	aat	agt	ggc	ggc	tat	ttt	gat	atg	aga	aat	att	1183
Leu	Tyr	Asn	Ala	Ser	Asn	Ser	Gly	Gly	Tyr	Phe	Asp	Met	Arg	Asn	Ile	
			300					305					310			
tta	aat	ggc	tct	gtc	gta	caa	aaa	cac	cct	ata	cat	gca	gtc	aca	ttt	1231
Leu	Asn	Gly	Ser	Val	Val	Gln	Lys	His	Pro	Ile	His	Ala	Val	Thr	Phe	
		315					320					325				
gtt	gat	aac	cat	gac	tct	cag	cca	gga	gaa	gca	ttg	gaa	tcc	ttt	gtt	1279
Val	Asp	Asn	His	Asp	Ser	Gln	Pro	Gly	Glu	Ala	Leu	Glu	Ser	Phe	Val	
	330					335					340					
caa	tcg	tgg	ttc	aaa	cca	ctg	gca	tat	gca	ttg	att	ctg	aca	agg	gag	1327
Gln	Ser	Trp	Phe	Lys	Pro	Leu	Ala	Tyr	Ala	Leu	Ile	Leu	Thr	Arg	Glu	
345					350					355					360	
caa	ggc	tac	cct	tcc	gta	ttt	tac	ggc	gat	tac	tac	ggc	ata	cca	act	1375
Gln	Gly	Tyr	Pro	Ser	Val	Phe	Tyr	Gly	Asp	Tyr	Tyr	Gly	Ile	Pro	Thr	
				365					370					375		
cat	ggc	gtt	cct	tcg	atg	aaa	tct	aaa	att	gat	cca	ctt	ctg	cag	gca	1423
His	Gly	Val	Pro	Ser	Met	Lys	Ser	Lys	Ile	Asp	Pro	Leu	Leu	Gln	Ala	
			380					385					390			
cgt	caa	acg	tat	gcc	tac	gga	acc	caa	cat	gat	tat	ttt	gat	cat	cat	1471
Arg	Gln	Thr	Tyr	Ala	Tyr	Gly	Thr	Gln	His	Asp	Tyr	Phe	Asp	His	His	
		395				400						405				
gat	att	atc	ggc	tgg	acg	aga	gaa	ggg	gac	agc	tcc	cac	cca	aat	tca	1519
Asp	Ile	Ile	Gly	Trp	Thr	Arg	Glu	Gly	Asp	Ser	Ser	His	Pro	Asn	Ser	
	410					415					420					
gga	ctt	gca	act	att	atg	tcc	gat	ggg	cca	ggg	ggc	aat	aaa	tgg	atg	1567
Gly	Leu	Ala	Thr	Ile	Met	Ser	Asp	Gly	Pro	Gly	Gly	Asn	Lys	Trp	Met	
425					430					435					440	

cat gtc ggg aaa cat aaa gct ggc caa gta tgg aga gat atc acc gga	1615
Pyr Val Gly Lys His Lys Ala Gly Gln Val Trp Arg Asp Ile Thr Gly	
445 450 455	
aat agg tct ggt acc gtc acc att aat gca gat ggt tgg ggg aat ttc	1663
Asn Arg Ser Gly Thr Val Thr Ile Asn Ala Asp Gly Trp Gly Asn Phe	
460 465 470	
act gta aac gga ggg gca gtt tcg gtt tgg gtg aag caa taaataagga	1712
Thr Val Asn Gly Gly Ala Val Ser Val Trp Val Lys Gln	
475 480 485	
acaagaggcg aaaattactt tcctacatgc agagctttcc gatcactcat acaccaata	1772
taaattggaa gctt	1786

<210> 2

<211> 516

<212> PRT

<213> Bacillus sp. KSM-AP1378

<400> 2

Met Lys Leu His Asn Arg Ile Ile Ser Val Leu Leu Thr Leu Leu Leu
-30 -25 -20

Ala Val Ala Val Leu Phe Pro Tyr Met Thr Glu Pro Ala Gln Ala His
-15 -10 -5 -1 1

His Asn Gly Thr Asn Gly Thr Met Met Gln Tyr Phe Glu Trp His Leu
5 10 15

Pro Asn Asp Gly Asn His Trp Asn Arg Leu Arg Asp Asp Ala Ala Asn
20 25 30

Leu Lys Ser Lys Gly Ile Thr Ala Val Trp Ile Pro Pro Ala Trp Lys
35 40 45

Gly Thr Ser Gln Asn Asp Val Gly Tyr Gly Ala Tyr Asp Leu Tyr Asp
50 55 60 65

Leu Gly Glu Phe Asn Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly Thr  
70 75 80

Arg Ser Gln Leu Gln Gly Ala Val Thr Ser Leu Lys Asn Asn Gly Ile  
85 90 95

Gln Val Tyr Gly Asp Val Val Met Asn His Lys Gly Gly Ala Asp Gly  
100 105 110

Thr Glu Met Val Asn Ala Val Glu Val Asn Arg Ser Asn Arg Asn Gln  
115 120 125

Glu Ile Ser Gly Glu Tyr Thr Ile Glu Ala Trp Thr Lys Phe Asp Phe  
130 135 140 145

Pro Gly Arg Gly Asn Thr His Ser Asn Phe Lys Trp Arg Trp Tyr His  
150 155 160

Phe Asp Gly Thr Asp Trp Asp Gln Ser Arg Gln Leu Gln Asn Lys Ile  
165 170 175

Tyr Lys Phe Arg Gly Thr Gly Lys Ala Trp Asp Trp Glu Val Asp Ile  
180 185 190

Glu Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Ile Asp Met Asp  
195 200 205

His Pro Glu Val Ile Asn Glu Leu Arg Asn Trp Gly Val Trp Tyr Thr  
210 215 220 225

Asn Thr Leu Asn Leu Asp Gly Phe Arg Ile Asp Ala Val Lys His Ile  
230 235 240

Lys Tyr Ser Tyr Thr Arg Asp Trp Leu Thr His Val Arg Asn Thr Thr  
245 250 255

Gly Lys Pro Met Phe Ala Val Ala Glu Phe Trp Lys Asn Asp Leu Ala  
260 265 270

Ala Ile Glu Asn Tyr Leu Asn Lys Thr Ser Trp Asn His Ser Val Phe  
275 280 285

Asp Val Pro Leu His Tyr Asn Leu Tyr Asn Ala Ser Asn Ser Gly Gly  
290 295 300 305

Tyr Phe Asp Met Arg Asn Ile Leu Asn Gly Ser Val Val Gln Lys His  
310 315 320

Pro Ile His Ala Val Thr Phe Val Asp Asn His Asp Ser Gln Pro Gly  
325 330 335

Glu Ala Leu Glu Ser Phe Val Gln Ser Trp Phe Lys Pro Leu Ala Tyr  
340 345 350

Ala Leu Ile Leu Thr Arg Glu Gln Gly Tyr Pro Ser Val Phe Tyr Gly  
355 360 365

Asp Tyr Tyr Gly Ile Pro Thr His Gly Val Pro Ser Met Lys Ser Lys  
370 375 380 385

Ile Asp Pro Leu Leu Gln Ala Arg Gln Thr Tyr Ala Tyr Gly Thr Gln  
390 395 400

His Asp Tyr Phe Asp His His Asp Ile Ile Gly Trp Thr Arg Glu Gly  
405 410 415

Asp Ser Ser His Pro Asn Ser Gly Leu Ala Thr Ile Met Ser Asp Gly  
420 425 430

Pro Gly Gly Asn Lys Trp Met Tyr Val Gly Lys His Lys Ala Gly Gln  
435 440 445

Val Trp Arg Asp Ile Thr Gly Asn Arg Ser Gly Thr Val Thr Ile Asn  
450 455 460 465

Ala Asp Gly Trp Gly Asn Phe Thr Val Asn Gly Gly Ala Val Ser Val  
470 475 480

Trp Val Lys Gln  
485

<210> 3

<211> 1753

<212> DNA

<213> Bacillus sp. KSM-K38

<220>

<221> sig\_peptide

<222> (162) .. (224)

<223>

<220>

<221> mat\_peptide

<222> (225) .. ()

<223>

<220>

<221> CDS

<222> (162) .. (1664)

<223>

<400> 3

gtatgcgaaa cgatgcgcaa aactgcgcaa ctactagcac tcttcaggga ctaaaccacc 60

ttttttccaa aaatgacatc atataaacia atttgtctac caatcactat ttaaagctgt 120

ttatgatata tgtaagcggt atcattaaaa ggaggtattt g atg aga aga tgg gta 176

Met Arg Arg Trp Val

-20



gta gca atg ttg gca gtg tta ttt tta ttt cct tcg gta gta gtt gca	224
Val Ala Met Leu Ala Val Leu Phe Leu Phe Pro Ser Val Val Val Ala	
-15 -10 -5 -1	
gat gga ttg aac ggt acg atg atg cag tat tat gag tgg cat ttg gaa	272
Asp Gly Leu Asn Gly Thr Met Met Gln Tyr Tyr Glu Trp His Leu Glu	
1 5 10 15	
aac gac ggg cag cat tgg aat cgg ttg cac gat gat gcc gca gct ttg	320
Asn Asp Gly Gln His Trp Asn Arg Leu His Asp Asp Ala Ala Ala Leu	
20 25 30	
agt gat gct ggt att aca gct att tgg att ccg cca gcc tac aaa ggt	368
Ser Asp Ala Gly Ile Thr Ala Ile Trp Ile Pro Pro Ala Tyr Lys Gly	
35 40 45	
aat agt cag gcg gat gtt ggg tac ggt gca tac gat ctt tat gat tta	416
Asn Ser Gln Ala Asp Val Gly Tyr Gly Ala Tyr Asp Leu Tyr Asp Leu	
50 55 60	
gga gag ttc aat caa aag ggt act gtt cga acg aaa tac gga act aag	464
Gly Glu Phe Asn Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly Thr Lys	
65 70 75 80	
gca cag ctt gaa cga gct att ggg tcc ctt aaa tct aat gat atc aat	512
Ala Gln Leu Glu Arg Ala Ile Gly Ser Leu Lys Ser Asn Asp Ile Asn	
85 90 95	
gta tac gga gat gtc gtg atg aat cat aaa atg gga gct gat ttt acg	560
Val Tyr Gly Asp Val Val Met Asn His Lys Met Gly Ala Asp Phe Thr	
100 105 110	
gag gca gtg caa gct gtt caa gta aat cca acg aat cgt tgg cag gat	608
Glu Ala Val Gln Ala Val Gln Val Asn Pro Thr Asn Arg Trp Gln Asp	
115 120 125	
att tca ggt gcc tac acg att gat gcg tgg acg ggt ttc gac ttt tca	656
Ile Ser Gly Ala Tyr Thr Ile Asp Ala Trp Thr Gly Phe Asp Phe Ser	
130 135 140	
ggg cgt aac aac gcc tat tca gat ttt aag tgg aga tgg ttc cat ttt	704
Gly Arg Asn Asn Ala Tyr Ser Asp Phe Lys Trp Arg Trp Phe His Phe	
145 150 155 160	
aat ggt gtt gac tgg gat cag cgc tat caa gaa aat cat att ttc cgc	752
Asn Gly Val Asp Trp Asp Gln Arg Tyr Gln Glu Asn His Ile Phe Arg	
165 170 175	
ttt gca aat acg aac tgg aac tgg cga gtg gat gaa gag aac ggt aat	800
Phe Ala Asn Thr Asn Trp Asn Trp Arg Val Asp Glu Glu Asn Gly Asn	
180 185 190	

tat gat tac ctg tta gga tcg aat atc gac ttt agt cat cca gaa gta	848
Tyr Asp Tyr Leu Leu Gly Ser Asn Ile Asp Phe Ser His Pro Glu Val	
195 200 205	
caa gat gag ttg aag gat tgg ggt agc tgg ttt acc gat gag tta gat	896
Gln Asp Glu Leu Lys Asp Trp Gly Ser Trp Phe Thr Asp Glu Leu Asp	
210 215 220	
ttg gat ggt tat cgt tta gat gct att aaa cat att cca ttc tgg tat	944
Leu Asp Gly Tyr Arg Leu Asp Ala Ile Lys His Ile Pro Phe Trp Tyr	
225 230 235 240	
aca tct gat tgg gtt cgg cat cag cgc aac gaa gca gat caa gat tta	992
Thr Ser Asp Trp Val Arg His Gln Arg Asn Glu Ala Asp Gln Asp Leu	
245 250 255	
ttt gtc gta ggg gaa tat tgg aag gat gac gta ggt gct ctc gaa ttt	1040
Phe Val Val Gly Glu Tyr Trp Lys Asp Asp Val Gly Ala Leu Glu Phe	
260 265 270	
tat tta gat gaa atg aat tgg gag atg tct cta ttc gat gtt cca ctt	1088
Tyr Leu Asp Glu Met Asn Trp Glu Met Ser Leu Phe Asp Val Pro Leu	
275 280 285	
aat tat aat ttt tac cgg gct tca caa caa ggt gga agc tat gat atg	1136
Asn Tyr Asn Phe Tyr Arg Ala Ser Gln Gln Gly Gly Ser Tyr Asp Met	
290 295 300	
cgt aat att tta cga gga tct tta gta gaa gcg cat ccg atg cat gca	1184
Arg Asn Ile Leu Arg Gly Ser Leu Val Glu Ala His Pro Met His Ala	
305 310 315 320	
gtt acg ttt gtt gat aat cat gat act cag cca ggg gag tca tta gag	1232
Val Thr Phe Val Asp Asn His Asp Thr Gln Pro Gly Glu Ser Leu Glu	
325 330 335	
tca tgg gtt gct gat tgg ttt aag cca ctt gct tat gcg aca att ttg	1280
Ser Trp Val Ala Asp Trp Phe Lys Pro Leu Ala Tyr Ala Thr Ile Leu	
340 345 350	
acg cgt gaa ggt ggt tat cca aat gta ttt tac ggt gat tac tat ggg	1328
Thr Arg Glu Gly Gly Tyr Pro Asn Val Phe Tyr Gly Asp Tyr Tyr Gly	
355 360 365	
att cct aac gat aac att tca gct aaa aaa gat atg att gat gag ctg	1376
Ile Pro Asn Asp Asn Ile Ser Ala Lys Lys Asp Met Ile Asp Glu Leu	
370 375 380	
ctt gat gca cgt caa aat tac gca tat ggc acg cag cat gac tat ttt	1424
Leu Asp Ala Arg Gln Asn Tyr Ala Tyr Gly Thr Gln His Asp Tyr Phe	
385 390 395 400	

gat cat tgg gat gtt gta gga tgg act agg gaa gga tct tcc tcc aga	1472
Asp His Trp Asp Val Val Gly Trp Thr Arg Glu Gly Ser Ser Ser Arg	
405 410 415	
cct aat tca ggc ctt gcg act att atg tcg aat gga cct ggt ggt tcc	1520
Pro Asn Ser Gly Leu Ala Thr Ile Met Ser Asn Gly Pro Gly Gly Ser	
420 425 430	
aag tgg atg tat gta gga cgt cag aat gca gga caa aca tgg aca gat	1568
Lys Trp Met Tyr Val Gly Arg Gln Asn Ala Gly Gln Thr Trp Thr Asp	
435 440 445	
tta act ggt aat aac gga gcg tcc gtt aca att aat ggc gat gga tgg	1616
Leu Thr Gly Asn Asn Gly Ala Ser Val Thr Ile Asn Gly Asp Gly Trp	
450 455 460	
ggc gaa ttc ttt acg aat gga gga tct gta tcc gtg tac gtg aac caa	1664
Gly Glu Phe Phe Thr Asn Gly Gly Ser Val Ser Val Tyr Val Asn Gln	
465 470 475 480	
taacaaaaag ccttgagaag ggattcctcc ctaactcaag gctttcttta tgctgcttag	1724
cttaacgctt ctacgacttt gaagcttta	1753

<210> 4

<211> 501

<212> PRT

<213> Bacillus sp. KSM-K38

<400> 4

Met Arg Arg Trp Val Val Ala Met Leu Ala Val Leu Phe Leu Phe Pro
-20 -15 -10

Ser Val Val Val Ala Asp Gly Leu Asn Gly Thr Met Met Gln Tyr Tyr
-5 -1 1 5 10

Glu Trp His Leu Glu Asn Asp Gly Gln His Trp Asn Arg Leu His Asp
15 20 25

Asp Ala Ala Ala Leu Ser Asp Ala Gly Ile Thr Ala Ile Trp Ile Pro
30 35 40

Pro Ala Tyr Lys Gly Asn Ser Gln Ala Asp Val Gly Tyr Gly Ala Tyr  
45 50 55

Asp Leu Tyr Asp Leu Gly Glu Phe Asn Gln Lys Gly Thr Val Arg Thr  
60 65 70 75

Lys Tyr Gly Thr Lys Ala Gln Leu Glu Arg Ala Ile Gly Ser Leu Lys  
80 85 90

Ser Asn Asp Ile Asn Val Tyr Gly Asp Val Val Met Asn His Lys Met  
95 100 105

Gly Ala Asp Phe Thr Glu Ala Val Gln Ala Val Gln Val Asn Pro Thr  
110 115 120

Asn Arg Trp Gln Asp Ile Ser Gly Ala Tyr Thr Ile Asp Ala Trp Thr  
125 130 135

Gly Phe Asp Phe Ser Gly Arg Asn Asn Ala Tyr Ser Asp Phe Lys Trp  
140 145 150 155

Arg Trp Phe His Phe Asn Gly Val Asp Trp Asp Gln Arg Tyr Gln Glu  
160 165 170

Asn His Ile Phe Arg Phe Ala Asn Thr Asn Trp Asn Trp Arg Val Asp  
175 180 185

Glu Glu Asn Gly Asn Tyr Asp Tyr Leu Leu Gly Ser Asn Ile Asp Phe  
190 195 200

Ser His Pro Glu Val Gln Asp Glu Leu Lys Asp Trp Gly Ser Trp Phe  
205 210 215

Thr Asp Glu Leu Asp Leu Asp Gly Tyr Arg Leu Asp Ala Ile Lys His  
220 225 230 235

Ile Pro Phe Trp Tyr Thr Ser Asp Trp Val Arg His Gln Arg Asn Glu  
240 245 250

Ala Asp Gln Asp Leu Phe Val Val Gly Glu Tyr Trp Lys Asp Asp Val  
255 260 265

Gly Ala Leu Glu Phe Tyr Leu Asp Glu Met Asn Trp Glu Met Ser Leu  
270 275 280

Phe Asp Val Pro Leu Asn Tyr Asn Phe Tyr Arg Ala Ser Gln Gln Gly  
285 290 295

Gly Ser Tyr Asp Met Arg Asn Ile Leu Arg Gly Ser Leu Val Glu Ala  
300 305 310 315

His Pro Met His Ala Val Thr Phe Val Asp Asn His Asp Thr Gln Pro  
320 325 330

Gly Glu Ser Leu Glu Ser Trp Val Ala Asp Trp Phe Lys Pro Leu Ala  
335 340 345

Tyr Ala Thr Ile Leu Thr Arg Glu Gly Gly Tyr Pro Asn Val Phe Tyr  
350 355 360

Gly Asp Tyr Tyr Gly Ile Pro Asn Asp Asn Ile Ser Ala Lys Lys Asp  
365 370 375

Met Ile Asp Glu Leu Leu Asp Ala Arg Gln Asn Tyr Ala Tyr Gly Thr  
380 385 390 395

Gln His Asp Tyr Phe Asp His Trp Asp Val Val Gly Trp Thr Arg Glu  
400 405 410

Gly Ser Ser Ser Arg Pro Asn Ser Gly Leu Ala Thr Ile Met Ser Asn  
415 420 425

Gly Pro Gly Gly Ser Lys Trp Met Tyr Val Gly Arg Gln Asn Ala Gly  
430 435 440

Gln Thr Trp Thr Asp Leu Thr Gly Asn Asn Gly Ala Ser Val Thr Ile  
445 450 455

Asn Gly Asp Gly Trp Gly Glu Phe Phe Thr Asn Gly Gly Ser Val Ser  
 160 465 470 475

Val Tyr Val Asn Gln  
 480

<210> 5

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 5  
 gagtcgacca gcacaagccc atcataatgg

30

<210> 6

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 6  
 taaagcttca atttatattg g

21

<210> 7

<211> 40

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 7

gggtcgacca gcacaagccg atggattgaa cggtacgatg

40

<210> 8

<211> 29

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 8

taaagctttt gttattgggtt cacgtacac

29

<210> 9

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 9

ccagatctac ttaccatttt agagtca

27

<210> 10

<211> 34

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 10

atttgccaaa tgacgggcag cattggaatc gggt

34

<210> 11

<211> 34

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 11

aaccgattcc aatgctgccc gtcatttggc aaat

34

<210> 12

<211> 34

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 12

tttgaatggc atttgtcaaa tgacggggaa ccac

34

<210> 13

<211> 33

<212> DNA



<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 13

acaaggagtc agttggaagg tgccgtgaca tct

33

<210> 14

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 14

cgaaaccaag taatatcagg t

21

<210> 15

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 15

aataccatt ccgattttaa atggcgc

27

<210> 16

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 16

gattgggatc agtcatgyca gcttcagaac aaa

33

<210> 17

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 17

aaattcaccg gaaaggtatg ggactgggaa gta

33

<210> 18

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 18

tcatccagat gtaatcaatg

20

<210> 19

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 19

cttagaaatt ggggagaatg gtatacaaat aca

33

<210> 20

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 20

gtgaaacata ttaaatgcag ctatacgaga gat

33

<210> 21

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 21

aacaccacag gtaaagaaat gtttgcagtt gca

33

<210> 22

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 22

agaattttgg caaatgacc t

21

<210> 23

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 23

ttgctgcaat ccataactat tttaat

26

<210> 24

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 24

cttgctgcaa tcgaaagyta tttaaataaa aca

33

<210> 25

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 25

ggctatttttg atatggcaaa tatttttaa at ggt

33

<210> 26

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 26

tctgacaagg cagcaaggtt a

21

<210> 27

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 27

gatccacttc tggaagcacg tcaaacg

27

<210> 28

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 28

gggggtaata aaagaatgta tgtcggg

27

<210> 29

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 29

atgtatgtcg ggcgacataa agctgg

26

<210> 30

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 30  
gatggttggg gggatttcac tgtaa

25

<210> 31

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 31  
ttcactgtaa acgatggggc agtttcg

27

<210> 32

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 32  
ggtttggtg cagcaataaa t

21

<210> 33

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<220>

<221> misc\_feature

<222> (16)..(18)

<223> n = a, c, t, or g

<400> 33

tttgaatggc atttgnnnaa tgacgggaac cac

33

<210> 34

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<220>

<221> misc\_feature

<222> (16)..(18)

<223> n = a, c, t, or g

<400> 34

aaattcaccg gaaagnnntg ggactgggaa gta

33

<210> 35

<211> 33

<212> DNA

<213> Artificial Sequence



<220>

<223> Synthetic DNA

<220>

<221> misc\_feature

<222> (16)..(18)

<223> n = a, c, t, or g

<400> 35

gtgaaacata tttaaannnag ctatacgaga gat

33

<210> 36

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<220>

<221> misc\_feature

<222> (16)..(18)

<223> n = a, c, t, or g

<400> 36

cttgctgcaa tcgaannnta tttaaataaaa aca

33

<210> 37

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 37

gatggttggg gggaattcac tgtaa

25

<210> 38

<211> 39

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 38

ccaacgaatc gttggcaggt aatttcaggt gcctacacg

39

<210> 39

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 39

attgatgcgt ggacgagttt cgacttttca ggg

33

<210> 40

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 40

tttcgacttt ccagggcgta a

21

<210> 41

<211> 43

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 41

ggtgttgact gggatcagca atatcaagaa aatcatatatt tcc

43

<210> 42

<211> 42

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<220>

<221> misc\_feature

<222> (27)..(27)

<223> n = a, c, t, or g

<400> 42  
catattttcc gctttgcaaa tacggtntgg aacaggcgag tg

42

<210> 43

<211> 44

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 43  
aatatcgact ttagtcatcc agatgtacaa gatgagttga agga

44

<210> 44

<211> 43

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 44  
gacgtaggtg ctctcgaatc ttatttagat gaaatgaatt ggg

43

<210> 45

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 45

cgataacatt ccagctaaaa a

21

<210> 46

<211> 38

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 46

gacctggtgg ttccaagaga atgtatgtag gacgtcag

38

<210> 47

<211> 42

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 47

aatggcgatg gatggggcga tttctttacg aatggaggat ct

42

<210> 48

<211> 39

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<220>

<221> misc\_feature

<222> (19)..(21)

<223> n = a, c, t, or g

<400> 48  
ccaacgaatc gttggcagnn natttcaggt gcctacacg

39

<210> 49

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 49  
gttgactggg atgagcgcaa acaagaaaat cat

33

<210> 50

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 50  
ggatgaaga gttcggtaat tatga

25

<210> 51

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 51  
agtcattccag aggtcgtaga tgagttgaag gat

33